Discrete Mathematics 19 (1977) 23–45.

THE ENUMERATION OF ARRAYS AND A GENERALIZATION RELATED TO CONTINGENCY TABLES*

I.J. GOOD

Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA

J.F. CROOK

School of Business Administration, Winthrop College, Rock Hill, South Carolina 29730, USA

Received 12 April 1976
Revised 17 August 1976

New methods are described for enumerating arrays, and for a generalization that is required for a Bayesian analysis of contingency tables. Exact and approximate formulae are given, together with some number-theoretic results.

1. Introduction

An array is a rectangle of non-negative integers, or the natural generalization in more than two dimensions. We assume that an array is two-dimensional unless otherwise stated. We shall be concerned with the problem of enumerating arrays with assigned marginal totals and with a more general problem related to the Bayesian analysis of contingency tables. The enumeration of arrays has received a fair amount of attention; see, for example, MacMahon [17], Anand, Dumir, and Gupta [2], Carlitz [3, 4], Grimson [12], Gupta [13], Gupta and Nath [14], Abramson and Moser [1], Stein and Stein [22], Stanley [21], and Good [10], and work cited in those papers. We shall not be concerned with "symmetrical" arrays because, apart from the combinatorial interest, we are anxious to apply the results to contingency tables and these are not usually symmetrical, although some are, for example, Good [9, p. 55].

Denote the numbers of rows and columns of an array by r and s, respectively, and the row and column totals by $n_{i.}$ $(i = 1, 2, \dots, r)$ and $n_{.j}$ $(j = 1, 2, \dots, s)$. Let $\sum n_{i.} = \sum n_{.j} = N$ and let

$$A(k,(n_{i.}),(n_{.j})) = \mathscr{C}\left(\prod_i x_i^{n_{i.}} \prod_j y_j^{n_{.j}}\right) \prod_{i,j} (1 - x_i y_j)^{-k} \qquad (k > 0), \tag{1.1}$$

* This work was supported in part by H.E.W., N.I.H. Grant #R01–GM18870–4. The theory is due to I.J. Good and the programming required for the application of Section 10 is due to J.F. Crook.

where $\mathscr{C}(\cdots)$ denotes "the coefficient of ... in". (The notion $(n_{i.})$, with parentheses, denotes the sequence of $n_{i.}$'s.) Then $A(1,(n_{i.}),(n_{.j}))$ is equal to the number of $r \times s$ arrays having marginal totals $(n_{i.})$ and $(n_{.j})$, whereas $A(k,(n_{i.}),(n_{.j}))$ is a more general expression whose value is required in a certain Bayesian approach to the analysis of contingency tables, Good [9, p. 52; 10]. The case $k = -1$, like $k = 1$, is also of combinatorial interest, since $(-1)^N A(-1,(n_{i.}),(n_{.j})) = A^*((n_{i.}),(n_{.j}))$, the number of arrays (incidence matrices) when all cell entries are 0 or 1.

Most of the literature on arrays is concerned with finding exact values for $A(1,(n_{i.}),(n_{.j}))$ in various special circumstances, for example, when $r = s$ and all the marginal totals are equal to n, so that we are dealing with "magical arrays" so to speak (since they are not quite magic squares), in which case we write $A(k,(n_{i.}),(n_{.j})) = A(k,n,r\times r)$. MacMahon [17, Vol. II, p. 161] was the first to prove that

$$A(1,n,3\times 3) = \binom{n+2}{2} + 3\binom{n+3}{4}, \tag{1.2}$$

and it was conjectured by Gupta [2, p. 768] that $A(1,n,r\times r)$ is of the form

$$A(1,n,r\times r) = \sum_{i=0}^{\binom{r-1}{2}} a_{r,i}\binom{n+r-1+i}{r-1+2i}. \tag{1.3}$$

Abramson and Moser [1] proved, using heavy algebra, and Smith [20] reported (but did not publish his proof) that

$$A(1,n,4\times 4) = \binom{n+3}{3} + 20\binom{n+4}{5} + 152\binom{n+5}{7} + 352\binom{n+6}{9}. \tag{1.4}$$

Stein and Stein [22] assumed Gupta's conjecture and programmed a "branching algorithm" on a computer specially designed for combinatorial calculations. It enabled them to calculate $A(1,n,r\times r)$ for $r = 4,5,6$ and for enough values of n to determine all the corresponding coefficients $a_{r,i}$ (together with an additional value of n to obtain a check of their calculations and of the conjecture). The conjecture was finally proved by Stanley [21].

Some other results can be inferred from Abramson and Moser [1]. For example, they generalized (1.2) by considering $r\times 3$ arrays with all row totals equal to n, and column totals $\mu, \nu, rn-\mu-\nu$ ($n \geq \mu$, $n \geq \nu$, $r \geq 2$), and they showed that

$$\begin{aligned} &A((n,n,n,\ldots,n),(\mu,\nu,rn-\mu-\nu)) = \\ &= \binom{\mu+r-1}{r-1}\binom{\nu+r-1}{r-1} - r\binom{\mu+\nu+2r-n-3}{2r-2}. \end{aligned} \tag{1.5}$$

A related special result is that if $n_{.1}+n_{.2}+\cdots+n_{.,s-1} \leq n_{i.}$ $(i = 1,2,\ldots,r)$, then the number of arrays is equal to the product from $j = 1$ to $j = s-1$ of the number of ordered partitions of $n_{.j}$ into r non-negative parts, that is,

$$A = \prod_{j=1}^{s-1} \binom{n_{\cdot j} + r - 1}{r - 1}. \tag{1.6}$$

The proof of this result is that corresponding to each of the $s - 1$ partitions there is clearly just one way of filling in the last column to satisfy the marginal totals.

Approximations and an asymptotic formula for $A(k,(n_{i\cdot}),(n_{\cdot j}))$ were given by Good [10], for "large" k. The first term of the asymptotic formula was found, in all the examples examined, to be within about 30% of the correct value even with k as small as 1 by Good [10] and Crook and Good [7], in other words for the number of arrays, but for the application to contingency tables values of k less than 1 are also of interest. In the present paper we give, among other things, (i) formulae for $A(k,2,r \times r)$, $A(k,n,2\times 2)$, $A(k,n,3\times 3)$ $(n \leq 4)$, and $A(k,3,r\times r)$ $(r \leq 4)$, (ii) new approximations for $A(k,(n_{i\cdot}),(n_{\cdot j}))$, (iii) some congruence properties for $A(1,n,r\times r)$ and for $A^*(n,r\times r)$, (iv) an algorithm for computing $A(k,(n_{i\cdot}),(n_{\cdot j}))$, and especially for $A(k,n,r\times r)$, using roots of unity, and a generalization to more dimensions; and (v) an outline of how the branching algorithm used by Stein and Stein [22] for the case $k = 1$ can be generalized.

2. Other formulae for $A(k,(n_{i\cdot}),(n_{\cdot j}))$

By expanding the various factors in the generating function for $A(k,(n_{i\cdot}),(n_{\cdot j}))$, given by (1.1), we see that

$$A(k,(n_{i\cdot}),(n_{\cdot j})) = \sum\nolimits^* \prod_{ij} \binom{m_{ij} + k - 1}{k - 1}, \tag{2.1}$$

where Σ^* denotes a summation over all arrays (m_{ij}) for which the marginal totals are $(n_{i\cdot})$ and $(n_{\cdot j})$. Thus $A(k,(n_{i\cdot}),(n_{\cdot j}))$ is a "weighted enumeration" of arrays, the weight associated with an array being $\Pi_{ij}\, w(m_{ij})$, where $w(m) = \binom{m+k-1}{m}$. Some of the analysis in this paper would be valid for any function w, and the right side of (1.1) would be replaced by

$$\mathscr{C}\left(\prod x_i^{n_{i\cdot}} \prod y_j^{n_{\cdot j}}\right) \prod_{ij} \sum_m w(m) x_i^m y_j^m. \tag{2.2}$$

When an array is a contingency table, the numbers m_{ij} denote frequencies in a sample. Formula (2.1) shows that the contribution to $A(k,(n_{i\cdot}),(n_{\cdot j}))$ from any one array depends on the frequencies of these frequencies; that is, if f_u of the rs frequencies m_{ij} are equal to u, then the contribution is

$$\prod \binom{u+k-1}{u}^{f_u} \quad (u = 0,1,2,3,\ldots), \tag{2.3}$$

which of course is always a *finite* product. Let $\mathcal{N}(f;(n_{i\cdot}),(n_{\cdot j}))$ denote the number of arrays having the marginal totals $(n_{i\cdot})$ and $(n_{\cdot j})$, and having the frequency count $f = (f_1, f_2, f_3, \ldots)$. (We omit f_0 from the notation because it does not affect the following formula.) Then

$$A(k,(n_{i.}),(n_{.j})) = \sum_{f} \mathcal{N}(f;(n_{i.}),(n_{.j})) \prod_{u} \binom{u+k-1}{u}^{f_u}. \tag{2.4}$$

When $k = -1$, we have $f_1 = N, f_2 = f_3 = \cdots = 0$, the product reduces to $(-1)^N$. and (2.4) states that $(-1)^N A(-1,(n_{i.}),(n_{.j})) = A^*((n_{i.}),(n_{.j}))$. $\mathcal{N}(f;(n_{i.}),(n_{.j}))$ may be described as "the frequency of the frequencies of the frequencies". When $k = 1$, (2.4) is obvious.

3. Formulae for $A(k,2,r \times r)$

Denote by b_r the number of $r \times r$ arrays with all marginal totals equal to 2 and no 2 in any cell, so that $b_r = A^*(2,r \times r) = A(-1,2,r \times r)$. All the entries in the array are 0's and 1's and, for example, $b_1 = 0$, $b_2 = 1$, and $b_3 = 6$. We define $b_0 = 1$. Then we have the recurrence relation for $r \geq 2$:

$$b_r = \binom{r}{2}[2b_{r-1} + (r-1)b_{r-2}], \tag{3.1}$$

which gives $b_4 = 90, b_5 = 2040, \ldots$. This relation is somewhat simpler than that published by Gupta et al. [2, p. 763], namely

$$b_r = \tfrac{1}{2}r(r-1)^2[(2r-3)b_{r-2} + (r-2)^2 b_{r-3}].$$

The following proof of (3.1) is of some interest.

Imagine a circuit starting at a cell containing a 1 in the top row of the array, proceeding to the other 1 in that row, then down to the other 1 in the column so reached, then across to the other 1 in the row so reached, and so on until we return to the original cell. (We certainly cannot return to any cell before arriving at the original one because all the others on the path are already connected directly to two 1's.) There are

$$\tfrac{1}{2}r(r-1)^2(r-2)^2\cdots(r-a)^2$$

such circuits possible, consisting of $2a+2$ edges $(a = 1,2,\ldots)$, and, if the $a+1$ rows and $a+1$ columns containing the 1's in such a circuit are removed, there will be b_{r-a-1} ways of completing the array. This proves that

$$b_r = \tfrac{1}{2}r\{(r-1)^2 b_{r-2} + (r-1)^2(r-2)^2 b_{r-3} + (r-1)^2(r-2)^2(r-3)^2 b_{r-4} + \cdots\}.$$

Therefore

$$2b_r/r = (r-1)^2 b_{r-2} + 2b_{r-1}/(r-1),$$

which proves (3.1).

We can express $A(k,2,r \times r)$ in terms of the b's, by computing the frequencies of the frequencies of the frequencies. For there are $\binom{r}{s}r(r-1)\cdots(r-s+1)$ ways of placing exactly s 2's in cells, no two 2's being in the same row or column; and, subject to any such placing, there are b_{r-s} ways of completing the $r \times r$ array to make all marginal totals equal to 2. Thus, using "partition notation", we have

$$\mathcal{N}(1^{2(r-s)}2^s) = \binom{r}{s} r(r-1)\cdots(r-s+1)b_{r-s}.$$

Therefore, by (2.4), we have

$$A(k,2,r \times r) = \sum_{s=0}^{r} \frac{[r(r-1)\cdots(r-s+1)]^2}{s!} \binom{k+1}{2}^s k^{2(r-s)} b_{r-s}. \tag{3.2}$$

From (3.2) we can deduce the following double exponential generating function:

$$\sum_{r=0}^{\infty} A(k,2,r \times r)x^r/(r!)^2 = e^{\frac{1}{2}kx}(1-k^2x)^{-\frac{1}{2}} \tag{3.3}$$

which generalizes the special case of $k = 1$ given by Gupta et al. [2, p. 764] and which they describe as "suggested by the referee". To prove (3.3) observe that the left side is equal to

$$\sum_{r=0}^{\infty} \sum_{s}^{s \le r} \frac{[r(r-1)\cdots(r-s+1)]^2}{(r!)^2 s!} \binom{k+1}{2}^s k^{2(r-s)} b_{r-s} x^r$$

$$= \sum_{s=0}^{\infty} \binom{k+1}{2}^s \frac{1}{s!k^{2s}} \sum_{r}^{r \ge s} \frac{(k^2x)^r b_{r-s}}{[(r-s)!]^2}$$

$$= \sum_{s=0}^{\infty} \binom{k+1}{2}^s \frac{x^s}{s!} \sum_{t=0}^{\infty} \frac{(k^2x)^t b_t}{(t!)^2}.$$

But, according to [2, p. 764],

$$\sum b_t y^t/(t!)^2 = e^{-\frac{1}{2}y}(1-y)^{-\frac{1}{2}} \tag{3.4}$$

and (3.3) follows after a little simplification. (Incidentally (3.4) is the case $k = -1$ of (3.3).) From this result we see that

$$\begin{aligned} A(k,2,r \times r) &= (r!)^2 2^{-r} \sum_{s=0}^{r} \frac{k^{r+s}}{(r-s)!2^s} \binom{2s}{s} \\ &= r!2^{-r} \sum_{s=0}^{r} (2s-1)!! \binom{r}{s} k^{r+s} \end{aligned} \tag{3.5}$$

which is somewhat simpler than (3.2) and will be useful for comparing orthodox and Bayesian significance tests, for some contingency tables having small cell expectations. (Here $1!! = 0!! = (-1)!! = 1, 6!! = 6 \cdot 4 \cdot 2$, etc.) The case $k = 1$ of (3.5) is given by Stein and Stein [22, p. 3] and our more general formula can also be obtained by their method.

By differentiating (3.3), after taking logarithms, we can deduce the recurrence relation (valid for $r = 2, 3, 4, \ldots$):

$$a_r(k) = [1 + (2r-1)k]a_{r-1}(k) - 2(r-1)ka_{r-2}(k), \tag{3.6}$$

where $r!k^r a_r(k) = 2^r A(k,2,r\times r)$, $A(k,2,0\times 0)=1$, and $A(k,2,1\times 1) = k(k+1)/2$. This recurrence relation is especially convenient for tabulating $A(k,2,r\times r)$ for any fixed value of k. The cases $k=1$ and $k=-1$ of (3.6) yield slight transformations of equation (18) of [2] and of our equation (3.1). The sequence $a_r(1)$ $(r=0,1,2,\ldots)$ begins 1, 2, 6, 28, 188, 1656, 17992, 232016.

4. 2×2 magical arrays

The 2×2 magical arrays are characterized by $r=s=2$, with all the marginal totals equal to n. By formula (2.1) we have,

$$A(k,n,2\times 2) = \sum_{\nu=0}^{n} \binom{\nu+k-1}{k-1}^2 \binom{n-\nu+k-1}{k-1}^2. \tag{4.1}$$

We can therefore deduce a generating function which can be written in terms of binomial coefficients and also in terms of hypergeometric functions or Legendre functions, thus

$$\begin{aligned}\sum_{n=0}^{\infty} A(k,n,2\times 2)x^n &= \left[\sum_{n=0}^{\infty}\binom{n+k-1}{k-1}^2 x^n\right]^2 \\ &= [F(k,k;1;x)]^2 \\ &= (1-x)^{2-4k}[F(1-k,1-k;1;x)]^2 \\ &= (1-x)^{-2k}\left[P_{k-1}\left(\frac{1+x}{1-x}\right)\right]^2 \\ &= (1-x)^{-2k}\left[P_{-k}\left(\frac{1+x}{1-x}\right)\right]^2.\end{aligned} \tag{4.2}$$

For the transformations used here see Erdélyi et al. [8, pp. 64 and 125], Whittaker and Watson [23, p. 312], and Pólya and Szegö [19, p. 92]. For example,

$$\sum_{n=0}^{\infty} A(2,n,2\times 2)x^n = (1+x)^2(1-x)^{-6}.$$

5. 3×3 magical arrays

For 3×3 magical arrays we make use of the "syzygetic" basis discovered by MacMahon [17, p. 160]. We define a *basis* for magical arrays as a set of arrays such that (i) every linear combination of members of the set, with non-negative integer coefficients, is a magical array; (ii) every magical array is so expressible; and (iii) no basis element is so expressible in terms of the others. Clearly a basis for $r\times r$ magical arrays must include all $r\times r$ permutation matrices, that is, the $r!$ matrices each of which consists of r 1's and (r^2-r) 0's the 1's being in distinct rows and columns. MacMahon proved that for $r=3$ the basis consists of the six permutation matrices alone and is connected by a single "syzygy" (or linear dependence with integral coefficients). The syzygy is

$$\boldsymbol{A}_1 + \boldsymbol{A}_2 + \boldsymbol{A}_3 = \boldsymbol{B}_1 + \boldsymbol{B}_2 + \boldsymbol{B}_3 \tag{5.1}$$

where

$$\boldsymbol{A}_\nu = \{\delta^\nu_{i-j}\}, \qquad \boldsymbol{B}_\nu = \{\delta^\nu_{j-i}\} \qquad (\nu = 1, 2, 3) \tag{5.2}$$

where the subscripts $i - j$ and $j - i$ are taken modulo 3, and Kronecker's δ's are to be understood. Each $\boldsymbol{A}$ is a simple "circulix" and each $\boldsymbol{B}$ a "skew circulix".

Owing to the syzygy (5.1), the expression of a magical array as a linear combination of the basis elements is not unique, but a "canonical" representation can be constructed. In fact, if an array $\boldsymbol{C}$ is of the form

$$\alpha_1\boldsymbol{A}_1 + \alpha_2\boldsymbol{A}_2 + \alpha_3\boldsymbol{A}_3 + \beta_1\boldsymbol{B}_1 + \beta_2\boldsymbol{B}_2 + \beta_3\boldsymbol{B}_3$$

then, by using the syzygy, we can subtract $\min(\beta_1, \beta_2, \beta_3)$ from each of β_1, β_2 and β_3 and add it to each of α_1, α_2, and α_3. Thus we can force at least one of the three β's to vanish, and this will supply our canonical representation of $\boldsymbol{C}$.

We can categorize the canonical ways of obtaining the magical arrays having marginal totals n according to the partitions of n into six parts

$$n = \alpha_1 + \alpha_2 + \alpha_3 + \beta_1 + \beta_2 + \beta_3,$$

where one of the β's is zero. The order of the α's among themselves, and the order of the β's among themselves, can be changed without affecting the frequency count of the array, and also the α's can be interchanged with the β's when $\alpha_1\alpha_2\alpha_3 = 0$. The number of distinct sequences $(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3)$ obtained from given *sets* $(\alpha_1, \alpha_2, \alpha_3)$ and $(\beta_1, \beta_2, \beta_3)$ then appears as a numerical coefficient in $A(k, n, 3 \times 3)$; for example, the cases where $n = 3$ and one α is 2 and one β is 1, or one α is 1 and one β is 2, provide 18 arrays each with the frequency count $0^4 1^2 2^2 3^1$. To see that this is the frequency count and to obtain the factors by which these coefficients are multiplied it is necessary to hold in mind that no two A's have a 1 common, and no two B's; but that each A has one 1 in common with each B. We thus obtain, by using (2.3),

$$A(k, 0, 3 \times 3) = 1 \tag{5.3}$$

$$A(k, 1, 3 \times 3) = 6k^3 \tag{5.4}$$

$$A(k, 2, 3 \times 3) = 6k^6 + 9k^4 \binom{k+1}{2} + 6\binom{k+1}{2}^3 \tag{5.5}$$

$$A(k, 3, 3 \times 3) = 6\binom{k+2}{3}^3 + 12\binom{k+1}{2}^3 k^3 + 18\binom{k+2}{3}\binom{k+1}{2}^2 k^2 + 18\binom{k+1}{2}^2 k^5 + k^9 \tag{5.6}$$

$$\begin{aligned} A(k,4,3\times 3) = {} & 6\binom{k+3}{4}^3 + 12\binom{k+2}{3}^3 k^3 + 18\binom{k+3}{4}\binom{k+2}{3}^3 k^2 \\ & + 6\binom{k+1}{2}^6 + 9\binom{k+3}{4}\binom{k+1}{2}^4 + 3\binom{k+1}{2}^3 k^6 \\ & + 18\binom{k+2}{3}^2\binom{k+1}{2} k^4 + 36\binom{k+2}{3}\binom{k+1}{2}^3 k^3 \\ & + 3\binom{k+1}{2}^3 k^6 + 9\binom{k+1}{2}^4 k^4. \end{aligned} \tag{5.7}$$

It can be generally proved that $A(k,n,r\times r)$ is a polynomial in k of degree nr, in fact $A(k,(n_i),(n_{\cdot j}))$ is of degree N (see (1.1)).

For computing $A(1,n,3\times 3)$ the argument is simpler because the individual contributions to the answer depend only on the number of zeros among a,b,c,d,e,f. We can thus obtain

$$\begin{aligned} A(1,n,3\times 3) = {} & 3\binom{n-1}{4} + \left[\binom{3}{2} + \binom{3}{1}^2\right]\binom{n-1}{3} + \left[1 + \binom{3}{2}\binom{3}{1} + \binom{3}{1}\binom{3}{2}\right]\binom{n-1}{2} \\ & + \binom{6}{4}\binom{n-1}{1} + \binom{6}{5}\binom{n-1}{0} \\ = {} & \binom{n+5}{5} - \binom{n+2}{5} - 3\binom{n+3}{4} + \binom{n+2}{2}, \end{aligned}$$

which result was otherwise obtained by MacMahon.

This method appears to be impracticable for $r > 3$.

6. A formula for $A(k,3,r\times r)$

The case $k = 1$ of the following formula is given by Stein and Stein [22] and this more general formula can be proved by using the generalized homogeneous product sums $h_i(k)$, defined in Section B1 of Good [10], in place of the $h_i(1)$'s used by Stein and Stein.

$$A(k,3,r\times r) = \frac{r!\,k^r}{6^r}\sum_{\lambda,\mu,\nu}\binom{r}{\lambda,\mu,\nu}\frac{2^\lambda 3^\mu(\mu+3\nu)!\,k^{\mu+2\nu}}{\nu!\,6^\nu}, \tag{6.1}$$

where λ,μ,ν run through all non-negative integers for which $\lambda+\mu+\nu=r$. This formula can be used for the tabulation of $A(k,3,r\times r)$ just as it was used for the case $k=1$ by Stein and Stein [22]. When k is equated to -1, (6.1) reduces to $(-1)^r$ times the formula for $A^*(3,r\times r)$ which was also stated and tabulated by Stein and Stein, where A^* denotes, as in Section 1, the number of arrays when all entries are 0 or 1. From (6.1) one can readily obtain the formulae

$$A(k,3,1\times 1) = (2k+3k^2+k^3)/6 = k(k+1)(k+2)/3! \tag{6.2}$$

$$A(k,3,2\times 2) = (2k^2+6k^3+11k^4+12k^5+5k^6)/9 \tag{6.3}$$

$$A(k,3,3\times 3)=(4k^3+18k^4+60k^5+153k^6+300k^7+315k^8+140k^9)/18 \qquad (6.4)$$

and

$$A(k,3,4\times 4)=4(2k^4+12k^5+58k^6+234k^7+813k^8+2250k^9+4060k^{10}+4200k^{11}+1925k^{12})/27 \qquad (6.5)$$

of which (6.2) is obvious from (2.3); (6.3), and (6.4) can be seen to agree with (4.1) and (5.6); and (6.5) correctly reduces to the values (2008 and 24) given in the tables of Stein and Stein when $k=1$ and when $k=-1$.

By writing

$$(\mu+3\nu)!=\int_0^\infty e^{-t}t^{\mu+3\nu}\,dt$$

we can transform (6.1) formally, after some straightforward manipulation, into the (divergent) double exponential generating function

$$\sum A(k,3,r\times r)(-x)^r/(r!)^2 \underset{F}{=} e^{-kx/3}\int_0^\infty e^{-t-\frac{1}{2}tk^2x}J_0(\sqrt{(t^3k^3x)/3})\,dt \qquad (6.6)$$

where the symbol F indicates that an equation is formal, and J_0 denotes the Bessel function of the first kind and order zero. We can deduce that

$$\sum A(k,3,r\times r)(-x)^r/(r!)^2 \underset{F}{=} e^{-kx/3}\sum_{\nu=0}^{\infty}\frac{(3\nu)!(-k^3x)^\nu}{36^\nu(\nu!)^2(1+\frac{1}{2}k^2x)^{3\nu+1}} \qquad (6.6a)$$

of which the case $k=1$ is mentiond by Comtet [6, p. 125]. (6.6a) could be used for the automatic computation of $A(k,3,r\times r)$, with the help of a program that performs polynomial algebra. Comtet states that there is, for $A(1,3,r\times r)$, a linear recurrence relation of the sixth order with coefficients that are polynomials in r, but he retracted this statement in a private communication in February, 1975.

When $k>0$, the integral in (6.6) is convergent near $x=0$ only for real and non-negative values of x. It then gives the sum of a divergent series by a method strongly resembling Borel's method (see Hardy [15, pp. 83 and 182]).

By putting $k=-1$ and changing the sign of x we obtain

$$\sum_{r=0}^{\infty} A^*(3,r\times r)(-x)^r/(r!)^2 \underset{F}{=} e^{-x/3}\int_0^\infty e^{-t+\frac{1}{2}tx}J_0(\sqrt{(t^3x)/3})\,dt \qquad (6.7)$$

$$\underset{F}{=} e^{-x/3}\sum_{\nu=0}^{\infty}\frac{(3\nu)!(-x)^\nu}{36^\nu(\nu!)^2(1-\frac{1}{2}x)^{3\nu+1}}.$$

It is interesting that $\sum A(k,n,r\times r)(-x)^r/(r!)^2$ $(n=1,2;\ k>0)$ can be written in forms resembling (6.6), namely

$$e^{-kx}=\int_0^\infty e^{-t}J_0(2\sqrt{(tkx)})\,dt \qquad (6.8)$$

and, from (3.3),

$$\int_0^\infty e^{-t-\frac{1}{2}kx}J_0(tk\sqrt{x})\,dt, \qquad (6.9)$$

respectively; but there may be no corresponding formula for $n \geq 4$. Formulae resembling (6.1) can be developed for $n \geq 4$, but they are much more complicated. The results for $n = 4$, $k = \pm 1$ are given in [22].

7. Approximations

Approximations for $A(k, (n_{i.}), (n_{.j}))$ are suggested in [10], including the following one for the case $k = 1$, which arises out of a statistical interpretation:

$$A(1, (n_{i.}), (n_{.j})) \approx \prod \binom{n_{i.} + s - 1}{n_{i.}} \prod \binom{n_{.j} + r - 1}{n_{.j}} \Big/ \binom{N + rs - 1}{N}. \tag{7.1}$$

For example, the ratio of the right side to the correct value of $A(1, n, 5 \times 5)$ is 0.708 for $n = 2$, 0.721 for $n = 4$, 0.724 for $n = 6$, and is 0.726 for $n = 8$. The ratio for $A(1, 3, 15 \times 15)$ is 0.652. An analogous approximation to A^* is

$$A^*((n_{i.}), (n_{.j})) \approx \prod \binom{s}{n_{i.}} \prod \binom{r}{n_{.j}} \Big/ \binom{rs}{N}. \tag{7.2}$$

For example, the ratio to the correct value of $A^*(2, 5 \times 5)$ is 1.500, and for $A^*(3, 15 \times 15)$ the ratio is 1.598. But O'Neill [18] has stated that, if all the marginal totals are less than $(\log r)^{\frac{1}{4}-\varepsilon}$ $(\varepsilon > 0)$, in other words for sufficiently sparse square arrays, then the asymptotic approximation, for $r = s \to \infty$, is

$$A^*((n_{i.}), (n_{.j})) \sim \frac{N!}{\prod n_{i.}! \prod n_{.j}!} \exp\left\{-\frac{2}{N^2} \sum \binom{n_{i.}}{2} \sum \binom{n_{.j}}{2}\right\} \tag{7.3}$$

of which a special case is

$$A^*(n, r \times r) \sim N!(n!)^{-2r} e^{-\frac{1}{2}(n-1)^2}. \tag{7.4}$$

For the two tables just mentioned, the ratios of this asympotic approximation to the correct values are 1.053 and 1.115, so that (7.3) may often be more accurate than (7.2) even when the array is much less sparse than is required by O'Neil's condition.

Stein and Stein mention that Everett has shown that

$$A(2, r \times r)/A^*(2, r \times r) \to e \quad \text{as } r \to \infty. \tag{7.5}$$

Now it can be seen from (3.5) that

$$A(k, 2, r \times r) \sim (2r)!(k/2)^{2r} \sum_{s=0}^{r} \frac{1}{s!(2k)^s} \sim 2\sqrt{\pi r}(kr/e)^{2r} e^{1/(2k)}, \tag{7.6}$$

the approximation being good if $|4kr| \gg 1$. We can deduce (7.5) by putting $k = 1$ and -1 and then taking the ratio. The first form of (7.6) is much the more accurate form when $|kr|$ is not large and this form is less than $A(k, 2, r \times r)$ when $k > 0$. The sum can be written in terms of the incomplete Gamma function. If kr is small, $A(k, 2, r \times r) \approx r!(k/2)^r$.

It is interesting to note from Stein and Stein's tables that the natural logarithms of $A(n, r \times r)/A^*(n, r \times r)$, for $r = 15$ and $n = 2, 3, 4$, and 5 are 1.035, 4.094, 9.190 and 16.37, so it is natural to ask whether

$$A(n, r \times r)/A^*(n, r \times r) \to e^{(n-1)^2} \qquad \text{when } r \to \infty. \tag{7.7}$$

(When $n = 2$, the ratio is even better approximated by $\exp[2r/(2r-1)]$.) This result has been proved by Everett and Stein [8A]. Their proof is lengthy.

The formulae (7.3) to (7.7) would all be implied by the following more general formula:

$$A(k, (n_{i\cdot}), (n_{\cdot j}), r \times s) \approx \frac{N!k^N}{\prod n_{i\cdot}! \prod n_{\cdot j}!} \exp\left\{\frac{2}{N^2 k} \sum \binom{n_{i\cdot}}{2} \sum \binom{n_{\cdot j}}{2}\right\} \tag{7.8}$$

which appears to be valid for large r and s, if the table is "sparse" in the sense that all cell "expectations" $n_{i\cdot} n_{\cdot j}/N$ are small. (This condition is much less stringent than O'Neil's sparsity condition.) We shall give a somewhat heuristic proof of (7.8) in outline. We think of the array as a contingency table with statistically independent rows and columns, so that the probability of the interior (n_{ij}), given the marginal totals $(n_{i\cdot})$ and $(n_{\cdot j})$, is equal to the Fisher–Yates expression

$$\frac{\prod n_{i\cdot}! \prod n_{\cdot j}!}{N! \prod n_{ij}!}. \tag{7.9}$$

From this it follows by a familiar argument [24, p. 215] that the expected value of $n_{ij}^{(\nu)}$ (meaning $n_{ij}(n_{ij}-1)\cdots(n_{ij}-\nu+1)$) is equal to $n_{i\cdot}^{(\nu)} n_{\cdot j}^{(\nu)}/N^{(\nu)}$ ($\nu = 0, 1, 2, \ldots$). Therefore the factorial moment generating function of the random variable n_{ij} is equal to the *finite* sum

$$\sum_{\nu=0}^{\infty} \frac{t^\nu n_{i\cdot}^{(\nu)} n_{\cdot j}^{(\nu)}}{\nu! N^{(\nu)}} = \sum_{\mu=0}^{\infty} p_\mu (1+t)^\mu,$$

where p_μ is the probability that $n_{ij} = \mu$. On putting $1 + t = u$, and equating coefficients of like powers of u on the two sides of the equation, we find that

$$p_\mu = \sum_{\nu=\mu}^{\infty} \frac{(-1)^\nu}{\mu!(\nu-\mu)!} \frac{n_{i\cdot}^{(\nu)} n_{\cdot j}^{(\nu)}}{N^{(\nu)}}. \tag{7.10}$$

Therefore f_2 (the number of cell frequencies equal to 2) has expectation

$$\sum_\nu \frac{(-1)^\nu}{2!(\nu-2)!} \frac{\sum_i n_{i\cdot}^{(\nu)} \sum_j n_{\cdot j}^{(\nu)}}{N^{(\nu)}} \sim X, \tag{7.11}$$

where $2X = \sum n_{i\cdot}^{(2)} \sum n_{\cdot j}^{(2)}/N^2$, whereas $E(f_3) \to 0$. Now, by (7.9), $\prod n_{i\cdot}! \prod n_{\cdot j}!/[N!(2!)^\nu]$ multiplied by the number of tables having $f_2 = \nu$ is about $e^{-X} X^\nu/\nu!$ (assuming approximately a Poisson distribution). Therefore the number of such tables is approximately

$$\frac{N!2^{\nu}}{\prod n_{i.}!\prod n_{.j}!}\,\frac{e^{-X}X^{\nu}}{\nu!}\,.$$

Therefore, from (2.4), we have

$$A(k,(n_{i.}),(n_{.j}))\approx\frac{N!}{\prod n_{i.}!\prod n_{.j}!}\,e^{-X}\sum_{\nu=0}^{\infty}k^{N-2\nu}X^{\nu}(\nu!)^{-1}[k(k+1)]^{\nu}$$

and (7.8) follows easily.

A special case of (7.8) is

$$A(k,n,r\times r)\sim\frac{(nr)!k^{nr}}{(n!)^{2r}}\exp\left\{\frac{(n-1)^2}{2k}\right\}\quad\text{as } r\to\infty. \tag{7.12}$$

Some examples of the ratio of $A(k,n,r\times r)$ to this asymptotic value are given in Tables 1 and 2. From Table 2 we infer that the asymptotic formula (7.12) gives reasonable approximations for $n=3$ when $k\geqslant 0.5$ even if r is not large; but, for k as small as 0.1, r needs to exceed 100 if the error is to be less than 50%. A more precise asymptotic formula would be useful. For $n=2$ the results are better, but $A(k,2,r\times r)$ can be computed extremely quickly anyway, by the methods of Section 3.

Table 1. Values of the quotient of $A(k,2,r\times r)$ by the asymptotic expression (7.12). (See also (7.6) and the comments following it.)

$r\backslash k$	0.1	0.5	1
4	1.060	1.209	1.086
10	1.666	1.058	1.028

Table 2. Values of the quotient of $A(k,3,r\times r)$ by the asymptotic expression (7.12).

$r\backslash k$	0.1	0.5	1	2	3
2	0.0000112	0.560	0.974	1.076	1.075
3	0.0000970	0.652	0.957	1.041	1.045
4	0.000441	0.674	0.953	1.028	1.032
5	0.00127	0.697	0.958	1.022	1.026
10	0.00845	0.809	0.977	1.011	1.013
15	0.0195	0.865	0.985	1.0074	1.0083
20	0.0387	0.896	0.989	1.0055	1.0065

$k=0.1$; r	30	40	50	60	70	80	90	100
Quotient	0.0928	0.156	0.218	0.276	0.327	0.374	0.415	0.451

8. Number-theoretical results

We now prove two simple divisibility theorems. Apart from their number-theoretical interest, they can be used for correcting almost exact values of $A(1, n, r \times r)$ and of $A^*(n, r \times r)$.

Theorem 8.1. *If r is a prime, then*

$$A(1, n, r \times r) \equiv 0 \pmod{r} \quad \text{if } n \not\equiv 0 \pmod{r},$$
$$\equiv 1 \pmod{r} \quad \text{if } n \equiv 0 \pmod{r}.$$

The result applies to A^* also, if $n \leqslant r$.

Proof. Consider an $r \times r$ array for which all the marginal totals are n but not all rows are identical. Then, by rotating the rows among one another cyclically, as if the rows were the generators of a spinning horizontal cylinder, we can produce just r such arrays (since r is here assumed to be prime), so that all such arrays fall into subsets each containing r arrays. Remove all these arrays, and apply the same idea to the remaining arrays (each of which has all its rows identical) but using columns instead of rows. The number of arrays removed will again be a multiple of r. The only array left, if any, will have all its rows identical and all its columns identical. There is no such array if n is not a multiple of r, and just one if n is a multiple of r. This argument can be readily generalized to more than two dimensions.

Theorem 8.2. *Let p be a prime number satisfying the inequalities $n < p \leqslant r$. (r need not be prime.) Then $A(1, n, r \times r)$ and $A^*(n, r \times r)$ are both multiples of p.*

Proof. Consider an $r \times r$ array for which all row and column totals are equal to n, and think of its first p rows as the generators of a horizontal cyclinder. These p rows cannot be identical, for then some of the column totals would be at least p and this is ruled out because $p > n$. Therefore, if this cylinder is spun, it will produce p distinct arrays, since p is a prime. Hence all legal arrays fall into non-overlapping sets each containing p arrays. This proof applies to both $A(1, n, r \times r)$ and $A^*(n, r \times r)$, and even to arrays with arbitrary constraints on each row partition if the constraints are the same for each row.

Much stronger results must be true. For it can be seen from the exact values of $A(1, n, r \times r)$ tabulated by Stein and Stein [22] that, for example, $A(1, 5, 4 \times 4) = M(2^4 \cdot 3^6)$, $A(1, 5, 5 \times 5) - 1 = M(5^3)$, $A(1, 5, 6 \times 6) = M(6^3)$, $A(1, 5, 8 \times 8) = M(2^{10} \cdot 3^3 \cdot 7)$, $A(1, 5, 12 \times 12) = M(7 \cdot 11 \cdot 12^5)$, while $A^*(5, 12 \times 12) = M(2^{10} \cdot 3^8 \cdot$ [illegible] $\cdot 7 \cdot 11^2)$, where M here denotes "is a multiple of". It is intriguing that these numbers have so many small factors.

9. A lemma concerning homogeneous generating functions and roots of unity

In the next section a method using roots of unity will be given for computing $A(k,(n_{i.}),(n_{.j}))$. As a check it has been used for confirming the formula

$$A(1,n,5\times 5)=\binom{n+4}{4}+115\binom{n+5}{6}+5390\binom{n+6}{8}+101275\binom{n+7}{10}$$
$$+858650\binom{n+8}{12}+3309025\binom{n+9}{14}+4718075\binom{n+10}{16} \qquad (9.1)$$

which was given by Stein and Stein [22]. The method of the next section depends on the following lemma which is of independent interest:

Lemma 9.1. *Let $P(\mathbf{v})=P(y_1,y_2,\dots,y_s)$ be a homogeneous polynomial in $y_1,y_2,\dots,y_s$ of degree N, and let $m_1,m_2,\dots,m_s$ be non-negative integers whose sum is N. Then*

$$\mathscr{C}(\mathbf{y}^{\mathbf{m}})P(\mathbf{y})=\frac{1}{\prod t_j}\sum_{\mathbf{v}}\boldsymbol{\omega}^{(\mathbf{t}-\mathbf{m})\mathbf{v}}P(\omega_1^{v_1},\dots,\omega_s^{v_s}), \qquad (9.2)$$

where $\mathbf{a}^{\mathbf{b}}$ denotes $a_1^{b_1}a_2^{b_2}\cdots a_s^{b_s}$, $\mathbf{a}^{\mathbf{bc}}$ denotes $a_1^{b_1c_1}a_2^{b_2c_2}\cdots a_s^{b_sc_s}$, $t_j>m_j$ is an integer $(j=1,2,\dots,s)$, $\omega_j=\exp(2\pi i/t_j)$, and the summation is over $v_j=0,1,2,\dots,t_j-1$ $(j=1,2,\dots,s)$.

For example, one may take $t_j=m_j+1$ $(j=1,2,\dots,s)$ in which case $\boldsymbol{\omega}^{(\mathbf{t}-\mathbf{m})\mathbf{v}}$ reduces to $\boldsymbol{\omega}^{\mathbf{v}}$. (It might have been thought necessary to take $t_j>N$ but to do so would be very uneconomical.)

Proof. By using familiar properties of roots of unity we may see that the right side is equal to the sum of the coefficients of

$$y_1^{l_1t_1-(t_1-m_1)}y_2^{l_2t_2-(t_2-m_2)}\cdots y_s^{l_st_s-(t_s-m_s)}$$

in $P(\mathbf{y})$, where $l_1,l_2,\dots,l_s$ are strictly positive integers. (The *technique* here is also familiar, but the exploitation of homogeneity that we make next may be new.) Since $P(\mathbf{y})$ is homogeneous of degree N we must have $\sum_j[l_jt_j-(t_j-m_j)]=N$. But then

$$N=\sum m_j\leq\sum l_jm_j=N-\sum(l_j-1)(t_j-m_j)\leq N$$

with equality only if $l_1=l_2=\cdots=l_j=1$. Thus the sum of the coefficients mentioned reduces simply to the required coefficient of $\mathbf{y}^{\mathbf{m}}$.

10. A general formula for $A(k,(n_{i.}),(n_{.j}))$ in terms of roots of unity

For brevity we write $n_{i.}=m_i$, $n_{.j}=n_j$ which enables us to write $A(k,(n_{i.}),(n_{.j}))$ as $A(k,\mathbf{m},\mathbf{n})$. From (1.1) we see that

$$A(k,m,n)=\mathscr{C}(y^n)\prod_{i=1}^{r}\mathscr{C}(x^{m_i})\prod_{j=1}^{s}(1-xy_j)^{-k},\tag{10.1}$$

where it may be noted that x requires no subscript. Now

$$\prod_{i=1}^{r}\mathscr{C}(x^{m_i})\prod_{j=1}^{s}(1-xy_j)^{-k}$$

is a *homogeneous* polynomial in $y_1, y_2, \ldots, y_s$ of degree N. Therefore by the lemma of Section 9 we have

$$A(k,m,n)=\frac{1}{\prod t_j}\sum_{\nu}\omega^{(r-n)\nu}\prod_{i=1}^{r}C(k,(\omega_j^{\nu_j}),m_i),\tag{10.2}$$

where

$$C(k,(\omega_j^{\nu_j}),m_i)=\mathscr{C}(x^{m_i})\prod_{j=1}^{s}(1-x\omega_j^{\nu_j})^{-k}.$$

In particular, if we take $t_j = n_j + 1$, we have

$$A(k,m,n)=\frac{1}{\prod(n_j+1)}\sum_{\nu}\omega^{\nu}\prod_{i=1}^{r}C(k,(\omega_j^{\nu_j}),m_i).\tag{10.3}$$

The function C can be written in a few different ways:

$$\begin{aligned}C(k,(\omega_j^{\nu_j}),m_i)&=\sum_{\mu_1,\ldots,\mu_s}^{|\mu|=m_i}\prod_{j=1}^{s}\binom{\mu_i+k-1}{k-1}\omega_j^{\mu_i\nu_j}\\&=\mathscr{C}(x^{m_i})\prod_{j=1}^{s}\sum_{l=0}^{m_i}\binom{l+k-1}{k-1}\omega_j^{l\nu_j}x^l\\&=\frac{1}{m_ir-m_i+1}\sum_{\xi=0}^{(r-1)m_i}\chi^{-\xi m_i}\prod_{j=1}^{s}\sum_{l=0}^{m_i}\binom{l+k-1}{k-1}\omega_j^{l\nu_j}\chi^{l\xi}\end{aligned}\tag{10.4}$$

where $|\mu| = \mu_1+\mu_2+\cdots+\mu_s$, and $\chi = \exp\{2\pi\sqrt{-1}/[(r-1)m_i+1]\}$. Here again we can instead take $\chi = \exp(2\pi\sqrt{-1}/u_i)$ where u_i is any integer greater than $(r-1)m_i$, and sum over ξ from 0 to $u_i - 1$, and this change may be useful if we make use of a Fast Fourier Transform algorithm for which u_i has certain favored values (see, for example, Good [11]).

A further transformation of $C(k,(\omega_j^{\nu_j}),m_i)$ can be made by summing the finite series

$$\sum_{l=0}^{m_i}\binom{l+k-1}{k-1}z^l\tag{10.5}$$

and this is apt to be especially worthwhile when $k = 1$. This gives

$$C(1,(\omega_j^{\nu_j}),m_i)=\frac{1}{m_ir-m_i+1}\sum_{\xi=0}^{(r-1)m_i}\chi^{-\xi m_i}\prod_{j=1}^{s}\Omega(m_i,r,\xi,\omega_j^{\nu_j}),\tag{10.6}$$

where

$$\Omega(m_i, r, \xi, \omega_j^{\nu_j}) = \begin{cases} \dfrac{1-(\omega_j^{\nu_j}\chi^{\xi})^{m_i+1}}{1-\omega_j^{\nu_j}\chi^{\xi}} & \text{if } \omega^{\nu_j}\chi^{\xi} \neq 1, \\ m_i + 1 & \text{if } \omega_j^{\nu_j}\chi^{\xi} = 1. \end{cases} \tag{10.7}$$

If the t_j's are chosen to be equal, to say t, so that ω_j can be called ω, then $C(k, (\omega^{\nu_j}), m_i)$ depends only on the frequency count (signature) of (ν_j) when k, t, and m_i are given, so that an appreciable reduction in computation time is then available.

Another tranformation of $C(k, (\omega_j^{\nu_j}), m_i)$ is

$$C(k, (\omega_j^{\nu_j}), m_i) = \mathscr{C}(x^{m_i}) \exp\left\{-k \sum_j \log(1 - x\omega_j^{\nu_j})\right\}$$

$$= \sum_{\alpha} \frac{k^{\alpha_1+\alpha_2+\cdots} S_1^{\alpha_1} S_2^{\alpha_2} \cdots}{\alpha_1!\alpha_2!\alpha_3! \cdots 1^{\alpha_1}2^{\alpha_2}3^{\alpha_3} \cdots}, \tag{10.8}$$

where the number of terms is equal to the number of partitions of m_i and in fact $\alpha_1, \alpha_2, \ldots,$ run through all non-negative integer solutions of

$$\alpha_1 + 2\alpha_2 + 3\alpha_3 + \cdots = m_i,$$

and

$$S_1 = \sum_j \omega_j^{\nu_j}, \qquad S_2 = \sum_j \omega_j^{2\nu_j}, \ldots .$$

Formula (10.8) is convenient for expressing $A(k, \boldsymbol{m}, \boldsymbol{n})$ as a polynomial in k and this is useful when its value is required for several values of k. A convenient check is obtained by putting $\boldsymbol{\nu} = \boldsymbol{0}$ and $sk = \lambda$, which shows that

$$\sum_{\alpha} \frac{\lambda^{\alpha_1+\alpha_2+\cdots}}{\alpha_1!\alpha_2!\alpha_3! \cdots 1^{\alpha_1}2^{\alpha_2}3^{\alpha_3} \cdots} = \binom{\lambda + m_i - 1}{\lambda - 1}. \tag{10.9}$$

For $\lambda = 1$, this can be proved by noting that $(\alpha_1!\alpha_2!\alpha_3! \cdots 1^{\alpha_1}2^{\alpha_2} \cdots)^{-1}$ is the probability that a "flat-random" permutation of m_i objects has α_1 cycles of length 1, α_2 of length $2, \ldots,$ a fact that is equivalent to an enumeration formula due to Cauchy [5].

When m_i is independent of i, and is equal to say m, we have

$$A(k, \boldsymbol{m}, \boldsymbol{n}) = A(k, m, \boldsymbol{n}) = \frac{1}{\prod (n_j + 1)} \sum_{\boldsymbol{\nu}} \boldsymbol{\omega}^{\boldsymbol{\nu}} [C(k, (\omega_j^{\nu_j}), m)]^r. \tag{10.10}$$

If in addition n_j is independent of j, and if $s = r$, we obtain

$$A(k, n, r \times r) = \sum_{\boldsymbol{\nu}} \omega^{|\boldsymbol{\nu}|} \left[\frac{C(k, (\omega^{\nu_j}), n)}{n+1}\right]^r \tag{10.11}$$

where $\omega = \exp[2\pi\sqrt{-1}/(n+1)]$ and each of $\nu_1, \nu_2, \ldots, \nu_r$ runs from 0 to n. The terms in (10.11) depend only on k, n, r, ω, and the frequency count of $(\nu_1, \nu_2, \ldots, \nu_r)$. The number of terms with frequency count q_0 noughts, q_1 ones, q_2 twos, ... is

$r!(q_0!q_1!\cdots q_n!)$, so we have

$$A(k,n,r\times r)=\sum_{q_0,\ldots,q_n}^{q_1+\cdots+q_n=r}\left\{\frac{r!\omega^{q_1+2q_2+\cdots+nq_n}}{q_0!q_1!\cdots q_n!}\left[\frac{C_q(k,n)}{n+1}\right]^r\right\}, \tag{10.12}$$

where $C_q(k,n)$ is given by the sum occurring in (10.8) with $\alpha_1+2\alpha_2+3\alpha_3+\cdots=n$, and $S_\nu=q_0+q_1\omega^\nu+q_2\omega^{2\nu}+\cdots$. We used (10.12) for most of the numerical calculations. The factor $(n+1)^{-r}$ could of course be taken outside the summation in (10.12), but for some reason, our calculations work better with the factor inside.

An example of (10.11) is

$$k^r r! = A(k,1,r\times r) \tag{10.13}$$

$$=\frac{k^r}{2^r}\sum_\nu(-1)^{|\nu|}\Big(\sum_j(-1)^{\nu_j}\Big)^r$$

$$=\frac{k^r}{2^r}\left[r^r-\binom{r}{1}(r-2)^r+\binom{r}{2}(r-4)^r-\cdots\right]$$

which proves the identity

$$r!2^r=r^r-\binom{r}{1}(r-2)^r+\binom{r}{2}(r-4)^r-\cdots$$

where the series continues to the term with coefficient $\binom{r}{r}$. This identity is a special case of the binomial identity

$$(a+bn)^r-\binom{r}{1}(a+bn-b)^r+\binom{r}{2}(a+bn-2b)^r-\cdots=b^r r!$$

which can be readily proved by a familiar technique of the calculus of finite differences.

Again (with the appropriate definitions of $S_1, S_2, \ldots$),

$$A(k,2,r\times r)=\frac{k^r}{6^r}\sum_\nu^{0,1,2}\omega^{|\nu|}(kS_1^2+S_2)^r. \tag{10.14}$$

where $\omega=\exp(2\pi\sqrt{-1}/3)$; and

$$A(k,3,r\times r)=\frac{k^r}{24^r}\sum_\nu^{0,1,2,3}(\sqrt{-1})^{|\nu|}(k^2S_1^3+3kS_1S_2+2S_3)^r. \tag{10.15}$$

11. Three-dimensional arrays

Consider a $q\times r\times s$ three-dimensional array for which the one-dimensional marginal totals are $n_{h..}, n_{.i.}, n_{..j}$ ($h=1,2,\ldots,q$; $i=1,2,\ldots,r$; $j=1,2,\ldots,s$) which we also denote, for brevity, by l_h, m_i, and n_j. We denote $\sum l_h=\sum m_i=\sum n_j$ by N, and we define $A(k,(l_h),(m_i),(n_j))$ or $A(k,\boldsymbol{l},\boldsymbol{m},\boldsymbol{n})$ by

$$A(k,\boldsymbol{l},\boldsymbol{m},\boldsymbol{n})=\mathscr{C}(\boldsymbol{x}^{\boldsymbol{l}}\boldsymbol{y}^{\boldsymbol{m}}\boldsymbol{z}^{\boldsymbol{n}})\prod_{h,i,j}(1-x_hy_iz_j)^{-k} \tag{11.1}$$

so that $A(1, \boldsymbol{l}, \boldsymbol{m}, \boldsymbol{n})$ is the number of three-dimensional arrays having the given one-dimensional marginal totals. We have

$$A(k, \boldsymbol{l}, \boldsymbol{m}, \boldsymbol{n}) = \mathscr{C}(\boldsymbol{z}^{\boldsymbol{n}})\left\{\mathscr{C}(\boldsymbol{y}^{\boldsymbol{m}})\left[\mathscr{C}\left(\prod x_h^{l_h}\right)\prod_{h,i,j}(1 - x_h y_i z_j)^{-k}\right]\right\}$$

$$= \mathscr{C}(\boldsymbol{z}^{\boldsymbol{n}})\left\{\mathscr{C}(\boldsymbol{y}^{\boldsymbol{m}})\prod_h \mathscr{C}(x_h^{l_h})\prod_{i,j}(1 - x_h y_i z_j)^{-k}\right\}$$

[because the factors containing $x_1, x_2, \ldots, x_q$ can be separated]

$$= \mathscr{C}(\boldsymbol{z}^{\boldsymbol{n}})\left\{\mathscr{C}(\boldsymbol{y}^{\boldsymbol{m}})\prod_h \mathscr{C}(x^{l_h})\prod_{i,j}(1 - x y_i z_j)^{-k}\right\}$$

$$= \mathscr{C}(\boldsymbol{z}^{\boldsymbol{n}})\frac{1}{\prod t_i}\sum_{\boldsymbol{\mu}} \omega^{(\boldsymbol{t}-\boldsymbol{m})\boldsymbol{\mu}}\prod_h \mathscr{C}(x^{l_h})\prod_{i,j}(1 - x\omega_i^{\mu_i} z_j)^{-k},$$

where $t_i \geq m_i + 1$, $\omega_i = \exp(2\pi\sqrt{-1}/t_i)$, and μ_i runs from 0 to $t_i - 1$ $(i = 1, 2, \ldots, r)$. This follows from the Lemma, as in the two-dimensional case, because

$$\prod_h \mathscr{C}(x^{l_h})\prod_{i,j}(1 - x y_i z_j)^{-k}$$

is a homogeneous polynomial in the y's of degree N (when the z_j's are regarded as constants). We have now expressed $A(k, \boldsymbol{l}, \boldsymbol{m}, \boldsymbol{n})$ as the coefficient of $\boldsymbol{z}^{\boldsymbol{n}}$ in a homogeneous polynomial in the z's of degree N Therefore

$$A(k, \boldsymbol{l}, \boldsymbol{m}, \boldsymbol{n}) = \frac{1}{\prod t_i \prod u_j}\sum_{\boldsymbol{\mu},\boldsymbol{\nu}} \omega^{(\boldsymbol{t}-\boldsymbol{m})\boldsymbol{\mu}}\psi^{(\boldsymbol{u}-\boldsymbol{n})\boldsymbol{\nu}}\prod_{h=1}^{q} C(k, (\omega_i^{\mu_i}), (\psi_j^{\nu_j}), l_h) \quad (11.2)$$

where

$$C(k, (\omega_i^{\mu_i}), (\psi_j^{\nu_j}), l_h) = \mathscr{C}(x^{l_h})\prod_{i,j}(1 - x\omega_i^{\mu_i}\psi_j^{\nu_j})^{-k}, \quad (11.3)$$

$u_j \geq n_j + 1$ $(j = 1, 2, \ldots, s)$, $\psi_j = \exp(2\pi\sqrt{-1}/u_j)$, and ν_j runs from 0 to $u_j - 1$ $(j = 1, 2, \ldots, s)$. The coefficient of x^{l_h} can of course be expressed in the various ways described for two-dimensional tables, with trivial modifications.

When $l_h = m_i = n_j = n$ (independent of h, i, and j) we write $A(k, \boldsymbol{l}, \boldsymbol{m}, \boldsymbol{n})$ as $A(k, n, r \times r \times r)$. Then

$$A(k, n, r \times r \times r) = \frac{1}{(n+1)^{2r}}\sum_{\boldsymbol{\mu},\boldsymbol{\nu}}^{0,1,\ldots,n} \omega^{|\boldsymbol{\mu}|+|\boldsymbol{\nu}|}\left[\mathscr{C}(x^n)\prod_{i,j}(1 - x\omega^{\mu_i+\nu_j})^{-k}\right]^r$$

$$= \frac{1}{(n+1)^{2r}}\sum_{\boldsymbol{\mu},\boldsymbol{\nu}} \omega^{|\boldsymbol{\mu}|+|\boldsymbol{\nu}|}[C(k, (\omega^{\mu_i+\nu_j}), n)]^r, \quad (11.4)$$

where $\omega = \exp[2\pi\sqrt{-1}/(n+1)]$, and

$$C(k, \omega^{\mu_i+\nu_j}, n) = \sum_{\alpha_1, \alpha_2, \ldots}^{\alpha_1+2\alpha_2+\cdots=n} \frac{k^{\alpha_1+\alpha_2+\cdots} S_1^{\alpha_1} S_2^{\alpha_2}\cdots}{\alpha_1!\alpha_2!\cdots 1^{\alpha_1}2^{\alpha_2}3^{\alpha_3}\cdots} \quad (11.5)$$

where

$$S_\beta = \sum_i \omega^{\beta\mu_i} \sum_j \omega^{\beta\nu_j} \qquad (\beta = 1,2,3,\dots).$$

When $n = 1$, the summand in (11.4) factorizes into a function of μ times the same function of ν and leads, by (10.13), to the result $k^r(r!)^2$ which is easily proved directly. Another example is

$$A(k,2,r\times r\times r) = \frac{k^r}{18^r}\sum_{\mu,\nu}^{0,1,2}\omega^{|\mu|+|\nu|}\left[k\left(\sum\omega^{\mu_i}\sum\omega^{\nu_j}\right)^2 + \sum_i\omega^{2\mu_i}\sum\omega^{2\nu_j}\right]^r, \quad (11.6)$$

where here $\omega = \exp(2\pi\sqrt{-1}/3)$.

A short computer program, based on formula (11.4), leads to values $A(1,2,2\times 2\times 2) = 12$, and $A(1,1,3\times 3\times 3) = 36$, both of which are easy to think out independently.

12. The branching algorithm

Stein and Stein [22] describe an interesting "branching" algorithm for evaluating $A(1,(n_i),(n_j))$ and we shall describe a more general algorithm for $A(k,(n_i),(n_j))$. Although this algorithm is very similar to theirs we think it is worth spelling out because of the difficulty we had in understanding their exposition.

Definition 12.1. If the components of a vector (n_i) of non-negative integers are rearranged in non-decreasing order we describe the rearrangement as the *signature* of (n_i), following Stein and Stein's terminology. Signatures can be put, in a natural manner, into one-to-one correspondence with "frequency counts" which can be expressed in partition notation such as $3^2 2^1 1^4 0^2$ meaning "two 3's, one 2, four 1's, and two 0's", the corresponding signature being $(3,3,2,1,1,1,1,0,0)$. The *name* of this signature is $3^2 2^1 1^4 0^2$.

Definition 12.2. A *pseudo-tree* is a finite connected oriented linear graph such that (i) each vertex has some predecessors (possibly none) and some successors (possibly none). (ii) There is a "top" vertex with no predecessors, a bottom vertex with no successors, and all other vertices have at least one predecessor and one successor (the pseudo-tree hangs from its root). The root is in *generation* or level 0. (iii) All the arrows on the edges of the graph point downwards. Every path from the top to a given vertex is of the same "length" called the generation g of that vertex $(g = 0,1,2,\dots)$.

Thus a pseudo-tree resembles an ordinary oriented tree except that in a pseudo-tree a vertex can have more than one predecessor, and the pseudo-tree has the shape of a spindle or Humpty-Dumpty.

Definition 12.3. Associated with each edge (which connects one vertex directly to another) is a non-negative number called the *strength* of that edge. The *strength of a path* from the top to the bottom is defined as the product of the strengths of the edges that make up that path. The *strength of the pseudo-tree* is defined as the sum of the strengths of all possible paths from the top to the bottom.

Definition 12.4. Consider a pseudo-tree that has just $s+1$ generations, and with ν_j vertices in the jth generation ($j=0,1,\ldots,s$), where $\nu_0=\nu_s=1$. We shall define matrices $M_1, M_2, \ldots, M_s$, where M_g may be called the gth matrix of the pseudo-tree. We first arrange all the vertices in each row of the pseudo-tree in an arbitrary order from left to right, so that we can talk of the first, second, ... vertices in each generation. Then M_g has ν_{g-1} rows and ν_g columns, and its element in row a and column b is the weight of the edge that connects the ath vertex in generation $g-1$ to the bth vertex in generation g. If these two vertices are not connected this weight is interpreted as zero. Thus M_g depends on the arbitrary ordering of the vertices within generations, but only up to permutations of the rows and permutations of the columns of M_g.

Theorem 12.5. *The strength of a pseudo-tree is equal to the matrix product $M_1M_2M_3\cdots M_s$, which is a scalar. This scalar does not depend on the arbitrary orders within generations.*

The proof is fairly obvious and is omitted.

We shall now associate with a pair of vectors $(n_{i.})$ and $(n_{.j})$ (regarded as the row totals and column totals of an array), a pseudo-tree whose strength is equal to $A(k,(n_{i.}),(n_{.j}))$. The construction is not symmetrical with respect to the rows and columns of the array.

The top vertex of the pseudo-tree is labelled with the *vector* $(n_{.j})$ and the *signature* of $(n_{i.})$, and may also be regarded as corresponding to some $r\times s$ arrays. Each vertex in the gth generation is labelled with the $(s-g)$-vector $(n_{.1}, n_{.2}, \ldots, n_{.s-g})$ and with a possible signature of the r-vector of row totals when the $g-1$ right-hand columns of the array are omitted ($g=1,2,\ldots,s$). A vertex in generation g can be regarded as corresponding to some $r\times(s-g)$ arrays. We now have to define the strengths of the edges.

Let a be a vertex in generation g and b a vertex in generation $g+1$ and let their labels be

$$(n_{.1}, n_{.2}, \ldots, n_{.s-g}; \mu_1, \mu_2, \ldots, \mu_r) \qquad \left(\sum \mu_i = n_{.1}+\cdots+n_{.s-g}\right).$$

and

$$(n_{.1}, n_{.2}, \ldots, n_{.s-g-1}; \mu'_1, \mu'_2, \ldots, \mu'_r) \qquad \left(\sum \mu'_i = n_{.1}+\cdots+n_{.s-g-1}\right).$$

Let $(\nu_1,\ldots,\nu_r)$ be a sequence of non-negative integers such that $(\mu_1-\nu_1,$

$\mu_2 - \nu_2, \ldots, \mu_r - \nu_r$) is some permutation of $(\mu'_1, \mu'_2, \ldots, \mu'_r)$. Of course $\sum \nu_i = n_{\cdot g}$. We may think of $(\nu_1, \ldots, \nu_r)$ as a possible gth column of an array. Then the strength of the edge joining a to b is defined as the sum of the product

$$\prod_{i=1}^{r} w(\mu_i - \nu_i),$$

where

$$w(\lambda) = \binom{\lambda + k - 1}{\lambda},$$

and where the sum is taken over all possible vectors $(\nu_1, \nu_2, \ldots, \nu_r)$ with the property just mentioned. This completes the definition of the pseudo-tree and the strengths of its edges, and gives a sufficient description of the branching algorithm for computing $A(k, (n_{i\cdot}), (n_{\cdot j}))$. For numerical examples, with $k = \pm 1$, see [22].

If w is replaced by an arbitrary function, and if the weight of an array is defined as $\prod_{ij} w(n_{ij})$, then the above algorithm would give the general "weighted enumeration" of all arrays with given row and column totals. For example, if $w(\nu) = 1$ when $\nu = 0$ or an integral power of 2, and otherwise $w(\nu) = 0$, then the weighted enumeration is simply the enumeration of arrays with given marginal totals when every cell contains either a 0 or a power of 2. It is even possible to allow w to depend also on the marginal totals, *so the algorithm could be useful for the theory of a wider class of Bayesian models for contingency tables than those considered by Good* [10].

13. Numerical examples

We hope to report most of our numerical examples in a statistical periodical because of the application to contingency tables. For the present we recall that we have programs for computing $A(k, n, r \times r)$ by the method depending on roots of unity. Within about 10%, the running times for the calculation on an IBM 370/168, using "double precision" FORTRAN, that is, using fourteen significant figures (and giving results correct to twelve or thirteen places) were, for the calculations done,

$$\frac{1}{680}\binom{n+r}{r} p(n) e^{-0.078n} \text{ seconds},$$

where $p(n)$ is the number of partitions of n (for example, $p(5) = 7$). For example, it would take about 25 minutes to compute $A(1, 7, 14)$. "Single precision" arithmetic (seven places) cuts only about 30% from the running time and gives proportional errors of the order of 1/10,000. The branching program, applied to $A(1, 7, 14)$, would take several hours on the MANIAC at Los Alamos, according to [22, p. 13], but would give the exact answer. A comparison of the roots-of-unity program with the branching program is difficult to make because the IBM machine, combined with FORTRAN, is very different from the Los Alamos MANIAC combined with the language Madcap V, which were used by Stein and Stein.

For $k = 1$ we did the calculations for $n = 3(1)8$, $r = 5(1)8$ and, where they overlapped with the results in [22], the results agreed to 13 significant figures. We found that the deviations from the exact values (given in [22]) were much the same as the magnitudes of the calculated imaginary parts (which would be zero in exact calculations), as one might expect. This enables us to assert, for example, some new values of $A(1, n, r \times r)$ given in Table 3. The approximations (7.1) and (7.12) are also given in the table. In these examples (7.1) is much the better approximation.

Table 3. Some values of $A(1, n, r \times r)$ and asymptotic approximations. Multiplication by 10^{15}, for example, is denoted by (15).

n	r	$A(1, n, r \times r)$	(7.1)	(7.12)
6	7	5.562 418 293 8(15)	3.88(15)	3.75(16)
7	7	2.157 176 080 5(17)	1.51(17)	5.85(18)
8	7	5.945 968 652 3(18)	4.16(18)	1.04(21)
6	8	1.146 012 423 8(20)	7.89(19)	6.39(20)
7	8	1.359 070 741 9(22)	9.37(21)	2.69(23)
8	8	1.046 591 482 7(24)	7.21(23)	1.14(26)

References

[1] M. Abramson and W.O.J. Moser, Arrays with fixed row and column sums, Discrete Math. 6 (1973) 1–14.

[2] H. Anand, V.C. Dumir, and H. Gupta, A combinatorial distribution problem, Duke Math. J. 33 (1966) 757–769.

[3] L. Carlitz, Enumeration of symmetrical arrays, Duke Math. J. 33 (1966) 771–782; 38 (1971) 717–731.

[4] L. Carlitz, Enumeration of 3 × 3 arrays, Fibonacci Q. 10 (1972) 489–498.

[5] A.L. Cauchy, Exersice d'analyse et de physique mathématique, iii (1844) 173 (cited, for example, by Ledermann [16] 73).

[6] L. Comtet, Advanced Combinatorics (D. Reidel, Dordrecht, 1974).

[7] J.F. Crook and I.J. Good, unpublished numerical calculations (1975).

[8] A. Erdélyi, et al. (Bateman Manuscript Project) 1953 Higher Transcendental Functions, I (McGraw-Hill, NY, 1953).

[8A] C.J. Everett and P.R. Stein, The asymptotic number of integer stochastic matrices, Discrete Math. 1 (1971) 55–72.

[9] I.J. Good, The Estimation of Probabilities: an Essay on Modern Bayesian Methods (M.I.T. Press, Cambridge, MA, 1965).

[10] I.J. Good, On the application of symmetric Dirichlet distributions and their mixtures to contingency tables, Ann. Stat. 4 (1976) 1159–1189.

[11] I.J. Good, The relationship between two fast Fourier transforms, IEEE Trans. Comput. C20 (March 1971) 310–317.

[12] R.C. Grimson, Enumeration of symmetric arrays with different row sums, Rend. Sem. Mat. Univ. Padova 48 (1972) 105–112.

[13] H. Gupta, On the enumeration of symmetric matrices, Duke Math. J. 38 (1971) 709–710.

[14] H. Gupta and G.L. Nath, Enumeration of stochastic cubes, Notices of the Amer. Math. Soc. 19 (1972) A-568.

[15] G.H. Hardy, Divergent Series (University Press, Oxford, 1949).

[16] W. Ledermann, Introduction to the Theory of Finite Groups (Oliver and Boyd, Edinburgh and London; Interscience, NY, 1949).

[17] P.A. MacMahon, Combinatory Analysis, Vol. 2 (University Press, Cambridge; Chelsea, New York, 1916/1960).
[18] P.E. O'Neil, Asymptotics and random matrices with row-sum and column-sum restrictions, Bull. Amer. Math. Soc. 75 (1969) 1276–1282. (We have corrected a misprint in the title.)
[19] G. Pólya and G. Szegö, Aufgaben und Lehrsätze aus der Analysis, II (Springer-Verlag, Berlin; Dover, NY, 1925/1945).
[20] D.A. Smith, The number of 4 × 4 magic squares, Notices Amer. Math. Soc. 18 (1971) 90–91.
[21] R.P. Stanley, Linear homogeneous diophantine equations and magic labellings of graphs, Duke Math. J. 40 (1973) 607–632.
[22] M.L. Stein and P.R. Stein, Enumeration of stochastic matrices with integer elements, Los Alamos Scientific Laboratory Report LA-4434 (1970).
[23] E.T. Whittaker and G.N. Watson, A Course of Modern Analysis, 4th edition (University Press, Cambridge, 1935).
[24] S.S. Wilks, Mathematical Statistics (University Press, Princeton, 1946).